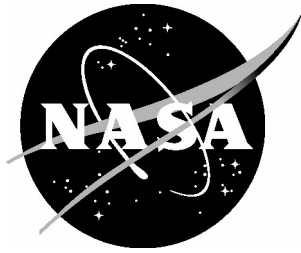


NASA/TM-2005-213505



Performance Comparison of Mainframe, Workstations, Clusters, and Desktop Computers

Douglas L. Farley
Langley Research Center, Hampton, Virginia

January 2005

The NASA STI Program Office . . . in Profile

Since its founding, NASA has been dedicated to the advancement of aeronautics and space science. The NASA Scientific and Technical Information (STI) Program Office plays a key part in helping NASA maintain this important role.

The NASA STI Program Office is operated by Langley Research Center, the lead center for NASA's scientific and technical information. The NASA STI Program Office provides access to the NASA STI Database, the largest collection of aeronautical and space science STI in the world. The Program Office is also NASA's institutional mechanism for disseminating the results of its research and development activities. These results are published by NASA in the NASA STI Report Series, which includes the following report types:

- **TECHNICAL PUBLICATION.** Reports of completed research or a major significant phase of research that present the results of NASA programs and include extensive data or theoretical analysis. Includes compilations of significant scientific and technical data and information deemed to be of continuing reference value. NASA counterpart of peer-reviewed formal professional papers, but having less stringent limitations on manuscript length and extent of graphic presentations.
- **TECHNICAL MEMORANDUM.** Scientific and technical findings that are preliminary or of specialized interest, e.g., quick release reports, working papers, and bibliographies that contain minimal annotation. Does not contain extensive analysis.
- **CONTRACTOR REPORT.** Scientific and technical findings by NASA-sponsored contractors and grantees.

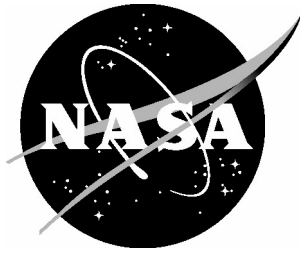
- **CONFERENCE PUBLICATION.** Collected papers from scientific and technical conferences, symposia, seminars, or other meetings sponsored or co-sponsored by NASA.
- **SPECIAL PUBLICATION.** Scientific, technical, or historical information from NASA programs, projects, and missions, often concerned with subjects having substantial public interest.
- **TECHNICAL TRANSLATION.** English-language translations of foreign scientific and technical material pertinent to NASA's mission.

Specialized services that complement the STI Program Office's diverse offerings include creating custom thesauri, building customized databases, organizing and publishing research results ... even providing videos.

For more information about the NASA STI Program Office, see the following:

- Access the NASA STI Program Home Page at [*http://www.sti.nasa.gov*](http://www.sti.nasa.gov)
- E-mail your question via the Internet to [*help@sti.nasa.gov*](mailto:help@sti.nasa.gov)
- Fax your question to the NASA STI Help Desk at (301) 621-0134
- Phone the NASA STI Help Desk at (301) 621-0390
- Write to:
NASA STI Help Desk
NASA Center for AeroSpace Information
7121 Standard Drive
Hanover, MD 21076-1320

NASA/TM-2005-213505



Performance Comparison of Mainframe, Workstations, Clusters, and Desktop Computers

Douglas L. Farley
Langley Research Center, Hampton, Virginia

National Aeronautics and
Space Administration

Langley Research Center
Hampton, Virginia 23681-2199

January 2005

The use of trademarks or names of manufacturers in the report is for accurate reporting and does not constitute an official endorsement, either expressed or implied, of such products or manufacturers by the National Aeronautics and Space Administration.

Available from:

NASA Center for AeroSpace Information (CASI)
7121 Standard Drive
Hanover, MD 21076-1320
(301) 621-0390

National Technical Information Service (NTIS)
5285 Port Royal Road
Springfield, VA 22161-2171
(703) 605-6000

Acronym and Abbreviation Table

AMD	Advanced Micro Devices
CFD	Computational Fluid Dynamics
CFL3D	Computational Fluids Laboratory 3-Dimensional flow solver
DDR	Double-Data-Rate Synchronous Dynamic Random Access Memory
DDR333	PC2700 Memory: 166 MHz actual clock rate, 333 MHz effective clock rate, 2.7 GB/s bandwidth per channel
Daemon	A class of computer program which runs in the background, rather than under the direct control of a user
DRAM	Dynamic Random Access Memory
GCC	GNU Compiler Collection – See http://gcc.gnu.org/
GB	Gigabyte
GHz	Gigahertz
GNU	“GNU's Not UNIX” – See http://www.gnu.org/
HP	Hewlett-Packard – See http://www.hp.com/
HPL	High Performance Linpack
IBM	International Business Machines – See http://www.ibm.com/
KB	Kilobyte
Kbyte	Kilobyte
L1 cache	A high speed, highest level cache in multilevel cache designs
L2 cache	The second, usually larger cache in multilevel cache designs
LAM-MPI	Implementation of the MPI Standard developed out of Indiana University in their Open Systems Laboratory – See http://www.lam-mpi.org/
LaRC	Langley Research Center
MB	Megabyte
MHz	Megahertz

MPI	Message Passing Interface – See http://www.mpi-forum.org/
MSI	Micro-Star Int’l Co.,Ltd
N	Column Height
NASA	National Aeronautics and Space Administration
NUMA	Non-Uniform Memory Access
OS	Operating System
P	When used in HPL, it is the row element to a HPL matrix
PC	Personal Computer
PGI	The Portland Group, Inc. The Portland Group is a wholly owned subsidiary of STMicroelectronics – See http://www.pgroup.com/
PoPC	Pile of Personal Computers
Q	When use in HPL, it is the column element to a HPL matrix
RAM	Random Access Memory
RPM	Rotations Per Minute
RDRAM	Rambus DRAM or RDRAM is a type of synchronous dynamic RAM, created by the Rambus Corporation
SGI	Silicon Graphics Incorporated – See http://www.sgi.com/
SMP	Symmetric Multi-Processor
SSE	Streaming SIMD Extensions
SSE2	Streaming SIMD Extensions Version 2
TB	Terabyte

Abstract

A performance evaluation of a variety of computers frequently found in a scientific or engineering research environment was conducted using a synthetic and an application program benchmark. From a performance perspective, emerging commodity processors such as the AMD Opteron have superior performance relative to legacy mainframe computers. The performance to price ratio of these new computers is substantially higher than the mainframe computers.

Clusters based upon pile-of-PCs design exhibited excellent performance as compared to the traditional mainframe computers. Newer PC processors like the Opteron were in excess of 130 percent faster than the available mainframe hardware. In many cases, the PC clusters exhibited comparable performance with traditional mainframe hardware when 8–12 processors were used. The main advantage of the PC clusters was related to their cost. Regardless of whether the clusters were built from new computers or whether they were created from retired computers their performance to cost ratio was superior to the legacy mainframe computers, with an 8 processor 2.4-GHz Pentium 4 cluster having twice the performance to cost of an 8 processor Silicon Graphics Incorporated (SGI) R14k.

Finally, the typical annual maintenance cost of legacy mainframe computers is several times the cost of new equipment such as multiprocessor Opteron workstations. The annual savings from eliminating the annual maintenance fee on legacy hardware can result in an annual increase in total computational capability for an organization.

Introduction

The phrase “if you build it, they will come” is frequently used to describe many situations in society and seems equally applicable to computers. As faster computers are developed, they facilitate the solution of more complex problems, which creates new technological and economic opportunities that encourage further computer advancements. Computation problems that in the 60’s and 70’s required room-size computers to solve are readily solved using commodity desktop computers today.

The performance trend of individual computers continues to follow Moore’s Law.¹ However, the largest gains in computer performance have not been solely through individual processor speed but have been through the clustering of computers into a massively parallel architecture called clusters. According to data from the Top 500 Supercomputer list, 40 percent of the fastest computers in the world are clusters of networked computers.²

A Beowulf cluster is “a kind of high-performance massively parallel computer built primarily out of commodity hardware components, running a free-software operating system like Linux or FreeBSD, and interconnected by a private high-speed network. It consists of a cluster of PCs or workstations dedicated to running high-performance computing tasks. The nodes in the cluster don’t sit on people’s desks; they are dedicated to running cluster jobs. It is usually connected to the outside world through only a single node.”³ A large number of these clusters are frequently composed of and referred to as a “Pile of

Personal Computers” (PoPC). A cluster designed around fault tolerance instead of speed is generically defined as a cluster and not specifically a Beowulf cluster.

In 1997, a team from Goddard Space Flight Center, headed by Dr. Donald J. Becker, achieved 1.25 Gflops (1 Gflop = 1 billion floating point operation per second) with 16-200 Mhz Pentium Pros.⁴ Using Dr. Becker’s original cluster as a model, 12.5 Gflops peak performance could be achieved with 16 of today’s 2-Ghz computers, neglecting general hardware and CPU architecture improvements. Similarly, a 16 processor-300Mhz-Silicon Graphics Incorporated Origin 2000 is capable of 9.6 Gflops,⁵ and 16 processor-675MHz-Cray T3E is capable of 24 Gflops.⁶ The SGI Origin would cost around \$300,000 and the Cray T3E would cost in excess of \$1,000,000 new, while the PC cluster would cost less than \$50,000.

Many research organizations, such as those at NASA’s Langley Research Center, have extensively used large multiprocessor machines, such as those produced by SGI, to meet their computational needs. Computers, such as SGI’s Origin series, can be configured with up to 512 processors, and 1 terabyte of RAM. The price of these computers is equally large. Newer hardware produced by Intel Corp. and Advanced Micro Devices (AMD) now are gaining competitive advantages against the traditional mainframe computers produced by manufacturers such as SGI, International Business Machines, and Hewlett-Packard/Compaq. In fact, Cray, a legacy mainframe computer manufacturer, is producing PoPC computer clusters using AMD’s new Opteron processor.

The objective of this study is to compare the performance of a variety of commonly used computer hardware employed in scientific and engineering applications (desktop PCs, SGI mainframe, PoPC clusters, and new desktop workstations). Both a synthetic benchmark (High Performance Linpack) and a scientific application program (Computational Fluids Laboratory 3-Dimensional flow solver) are used in this study. A discussion of the relative performance and cost of these machines running the different software is presented, which will aid scientists and engineers in future computer purchases.

Scope of Investigation

The objective of this study is to provide information that will aid the scientist or engineer to determine the type of computer resources necessary to meet their requirements. With the vast number of computers and software applications commonly used in a scientific or engineering environment, it is virtually impossible to include all of them in this study. Therefore, the scope of this study was limited to the test and production hardware and software applications available in our lab. This section consists of subsections describing the benchmark software and the computer hardware.

Benchmark Software

There were two benchmark software programs used in this study: a synthetic benchmark, High Performance Linpack (HPL); and a scientific application program, Computational Fluids Laboratory 3-Dimensional flow solver (CFL3D). Jack Dongarra and R. Clint Whaley of the Innovative Computing Laboratory at the University of Tennessee wrote HPL.⁷ HPL is the de facto benchmark of Beowulf and multiprocessor high performance computing systems, and is the official multiprocessor benchmark of the top 500-supercomputer list.⁸ HPL, more specifically Linpack, is a collection of Fortran and C subroutines that analyze and solve linear equations and linear least-squares problems to calculate the potential performance of a system. Linpack is a floating point intensive benchmark and reports speed in floating-point operations per second or FLOPS. A deficiency of High Performance Linpack is its method of measuring the performance of a heterogeneous cluster, that is, a cluster comprised of nonsimilar nodes.

As measured by HPL, a cluster's performance is proportional to the performance of the slowest processor multiplied by the number of processors in the cluster. It is important to remember this when comparing Linpack performance results from heterogeneous and homogeneous clusters.

The main input parameter to HPL is a block size, or column height (N), which can be converted into an average memory footprint used. For example, if you have 1 GB of memory (1073741824 bytes), and you assume eight bytes per element, you have approximately 130 million eight-byte elements. The maximum block size is the square root of the 130-million elements or a maximum N of 11585. The HPL configuration file HPL.dat, shown in figure 1, was configured the same on all systems except for variables P, Q,⁹ which are the row and column sizes used to subdivide the HPL test matrices, and the block size allocation when they had to be altered due to a machine's processor count and available memory.

The second benchmark program is CFL3D using the ONERA M6 wing test case. CFL3D, a NASA developed Navier-Stokes Computational Fluid Dynamics (CFD) 3-dimensional flow solver, is maintained by Dr. Robert T. Biedron of the Computational Modeling and Simulation Branch at NASA Langley Research Center (LaRC).¹⁰ The ONERA M6 wing data used in CFL3D were based upon the work by Schmitt and Charpin in 1979.¹¹ "This widely used test case consists of an isolated wing in a transonic free stream of Mach 0.84 at an angle of attack of 3.06° with a chord Reynolds number of 21.66 million."¹² The M6 wing is a standard test case for the CFL3D suite. An example of the surface grid for the M6 wing and the calculated pressure contours is seen in figure 2. CFL3D is used to solve steady and unsteady flows (including turbulence) using multigrid and mesh sequenced convergence acceleration options. CFL3D V6, used in this study, supports Message Passing Interface (MPI). The M6 test case run on CFL3D is representative of scientific and engineering application codes that have both floating point and disk I/O intensive functions. Other commercial software of this genre are structural analysis finite element computer codes such as NASTRAN¹³ and Abaqus.¹⁴

Computer Hardware, Operating System, and Environment

A wide variety of computers, representative of those frequently used in NASA LaRC scientific computing environments, was chosen for this study. These computers included three SGI machines, two AMD Opteron based workstations, three PC based clusters, and a dual processor desktop PC. A table representing the computational hardware used in this study and discussed in this section can be seen in table 1.

Three different SGI MIPS¹⁵ 64-bit processor based computers were included in this study as representative legacy computer systems: an 8-processor R14000¹⁶ computer (called South), an 8-processor R12000¹⁷ computer (Norm), and a 16-processor R10000¹⁸ computer (Whitcomb). South, a SGI Origin 300, has eight 600-MHz IP35 R14000 processors with 8 GB of Main memory, a 32-KB instruction cache, 32-KB data cache, and a 4-MB L2 cache. South when new (2001) cost approximately \$110K. Norm, a SGI Onyx 2, has eight 400-Mhz IP27 R12000 processors and has 7.5 GB of main memory, 32-KB data and instruction caches, and an 8-MB L2 cache. Initial cost in 1999 was in excess of \$300K. Whitcomb, an SGI Origin 2000, has sixteen 250-MHz IP27 R10000 processors with 32-KB data and instruction caches and a 4-MB L2 cache and also cost approximately \$300K in 1998. The annual maintenance fee for these SGI machines is on the order of \$40,000 each. The MIPS-pro 7.3 compilers were used along with GCC, for CFL3D and HPL, respectively. All three machines run IRIX 6.5 as the base OS.

Two AMD Opteron based workstations were included in this study. The dual processor Opteron 242 (1.6-GHz) computer (McCoy), has 4 GB of DDR333 RAM running on an MSI K8D Master Motherboard with one 80-GB 7200 rpm hard disk. This Opteron system was a state of the art commodity desktop workstation computer and in 2003 cost approximately \$4000. The second workstation is a quad-processor AMD Opteron 844 (1.8-GHz) computer (Eureka). It has 32 GB of DDR333 RAM running on a Newisys Opteron Quartet system, which runs the AMD8111+8131 chipsets. Both workstations use the SuSE Enterprise 7.1 operating system. Both included the PGI 5.0-1 compilers that are necessary for running CFL3D, GCC 3.3, and LAM-MPI 6.5.8. LAM-MPI was built for both 32- and 64-bit applications with both PGI and GCC, with compiler optimization level 2 and the 32- and 64-bit executable flags, respectively.

Three PC clusters were also evaluated in this study. The cluster Geowulf consists of eight 2.4-GHz Pentium processors with 1.5-GB RDRAM per processor. The nodes are connected across a 100BaseT Ethernet on an HP Procurve 2524 Fast-Ethernet switch. Geowulf runs a ROCKS 2.2 Linux cluster operating system and cost approximately \$20,000 in 2003. The second cluster, EABWulf, is a heterogeneous cluster consisting of 10 dual 550 to 866 Mhz Intel Xeon processors, a dual Xeon processor Master node, and a dual Xeon processor file server assembled from parts of retired desktop machines. Each node has 1 GB of RAM. All nodes run a modified Red Hat Linux 7.3 distribution and interconnected using a 100BaseT Ethernet on a 3com 12 Port SuperStack3 switch. The cost to assemble this cluster was the purchase of the switch and incidental hardware, requiring less than \$5,000 in new expenditures. The original cost of the computers for EABWulf would have been in excess of \$60K. The third cluster BWolf was also assembled from parts of retired machines. Its master node was a dual Athlon MP 1200, the eight compute nodes consisted of Intel Pentium II's and III's ranging from 400–700 Mhz, and each node had between 256 MB to 1 GB of RAM. These nodes were connected to a Cisco Catalyst 2900 switch, and all machines used the same modified Red Hat Linux 7.3 distribution as EABWulf. The total cost of this system was under \$2,500, since only the dual Athlon master node and a few network cards were new expenditures. The original cost of the BWolf compute nodes would have been in excess of \$20K. The expected performance-to-price of these two PoPC clusters is an excellent example of the potential use of computing resources that otherwise would be ignored.

Benchmarking individual desktop computers relative to mainframe and PC clusters is beneficial. A dual processor desktop computer (called Riker) is included in this comparison. Riker has dual AMD Athlon MP 1900 processors running on a Tyan S2462UNG motherboard with 2 GB of PC2100 memory. A modified Redhat Linux 7.3 operating system was used. Riker's cost in 2002 was approximately \$2,100.

The Message Passing Interface interconnects LAM-MPI¹⁹ was chosen. The LAM-MPI implementation of MPI and MPI2 was developed by Indiana University²⁰ but originated from the Ohio Supercomputer Center.²¹ LAM-MPI has evolved into a modular MPI development environment that spawns a user daemon on each compute node to handle the MPI runtime environment. It features the ability to clean up memory space of previous jobs, to integrate with several process schedulers, and to integrate with multiprocessor machines where it will pass data using shared memory communication. A preliminary study demonstrated that the added features of LAM-MPI, as well as its similar runtime to other MPI implementations made it a better choice for MPI data transport. LAM-MPI was built with GCC and the PGI compilers, HPL was built with GCC, and because of its dependencies, CFL3D was built with the PGI compilers.

Results and Discussion

The results are presented relative to the two benchmarks for each machine. Benchmarks were not run on all machines due to compilation errors that could not be resolved using the available compilers. An assessment of the performance relative to system cost is also included.

HPL Benchmark Results

Due to system configuration issues, the SGI computers were unable to properly execute HPL. Therefore, no HPL results are presented for the three SGI machines in this study. As previously mentioned HPL is a good benchmark for measuring the performance of homogeneous clusters. Therefore, EABWulf was reconfigured two ways, as four-single-processor nodes and four-dual-processor nodes with each processor operating at 866 MHz. Both of these cluster configurations are homogeneous; however, the eight-processor cluster is a symmetric multiprocessor (SMP) kernel with memory sharing. The performance as a function of block size of these two configurations of EABWulf is presented in figure 3. For a block size of 1000 elements the performance of these computers is nearly identical. As the block size increases the performance of both computers increases. The four-processor cluster's performance becomes constant at block sizes of 15000 elements whereas the eight-processor cluster's performance does not reach a constant performance at a block size of 20000 elements. Based upon data trends relative to the performance of the four-processor cluster it is reasonable to expect the eight-processor cluster to reach a maximum performance state for block sizes slightly greater than 20000 elements.

The performance of the four- and eight-processor clusters with a block size of 20000 elements is 1.21 and 2.24 Gflops, respectively. These results demonstrate that a SMP machine can produce scalable performance with sufficiently large block sizes. The lack of scalability of results for smaller block sizes is due to the amount of time associated with data transmission between nodes. These results prove that dual processor machines sharing their local memory can perform as fast or faster than two-single-processor machines, which can be an important issue in purchasing or developing a cluster. Small undulations in the performance curves are present for both four- and eight-processor machines. These undulations are insignificant in magnitude.

The two Processor AMD Opteron 242 (1.6 GHz), named McCoy, and the four Processor AMD Opteron 844 (1.8 GHz), named Eureka, were also benchmarked using HPL. The Opteron workstation, McCoy, was a preproduction computer that was on loan for evaluation. During this evaluation a commercial version of the PGI and AMD64 optimized versions of the GCC compiler were not available. Therefore, properly tuned 32-bit libraries for HPL could not be created for McCoy. The lack of compilers was later remedied for Eureka, the four-processor Opteron workstation.

HPL allows for specifying square or rectangular block geometry, by means of the P and Q values from figure 1. Performance differences between square and rectangular blocks were found to be small relative to the effects of block size, number of processors, and whether the program was compiled for 32- or 64-bit executable modes. Although analyses were conducted to investigate the effect of block geometry on performance, the results are not included in this study and it was decided to follow the general guidelines of the HPL team.²²

The two-processor workstation, McCoy, and four-processor workstation, Eureka, were run with HPL compiled in 64-bit executable mode. Performance results are presented in figure 4. Using a 15000 block size, a 2.51- and 3.06-Gflops performance, respectively, was achieved with one processor. For a block

size of 20000 elements, the performances were 4.67 and 5.96 Gflops, respectively, with two processors. The four-processor Opteron had a 27 percent higher performance than the two-processor Opteron even though the clock speed difference was only 12 percent. This disproportionate performance increase is attributed to architectural changes. Performance substantially increases with an increasing number of processors. Eureka, using two processors, achieved a 94-percent increase in performance relative to single processor performance, and a 250-percent increase in performance using four processors, as seen in figure 4. The increase in performance achieved using four processors was 80 percent relative to the two-processor performance. This performance increase was less than the 94-percent increase in performance achieved between the one to two processor results. The reduced increase in performance is attributed to the nonuniform memory access (NUMA)²³ architecture employed in the Opteron.

The HPL measured performance of Eureka in 32-bit executable mode was within a few percent of the speed achieved in 64-bit, as shown in figure 4. These results were initially surprising but upon closer examination should have been expected. The SSE/SSE2 floating point register size is 128 bits for both 32- and 64-bit operations; therefore, the same amount of time is required to perform the same operations, hence similar performance. There are, however, more available registers in the 64-bit operation, which could account for the approximately 0.25-Gflop increase from 32- to 64-bit runs. Although there is not an extensive speed difference between 32- and 64-bit HPL runs, the main advantage of the 64-bit operation is related to available address space, which can be important for certain classes of problems. Modern 64-bit processors can address up to 1 TB of memory creating a block size of 65000 elements whereas a 32-bit processor can address only up to 4 GB producing a block size of 23000 elements.

CFL3D Results

CFL3D is a scientific and engineering application program that has both floating point computational and I/O intensive functions. CFL3D was executed on all of the aforementioned computers except the quad-processor Opteron Eureka. Compiler related problems prohibited running CFL3D on Eureka and for the 64-bit dual-processor case on the Opteron McCoy. Wall clock execution time as a function of the number of processors for all computers is presented in figure 5. All of the hardware was tested in a dedicated or shared system mode, with similar low or negligible load levels and load averages on systems where it was not possible to get a dedicated mode before testing.

Total execution time significantly decreases for all computers as a function of increasing number of processors. The steepest slope of these data is between one and two processors, and the PC clusters and workstations exhibit a steeper slope than the SGI hardware, as shown in figure 5. The multiprocessor SGI machine, Whitcomb, and the PC cluster, EABWulf, achieve a constant performance state at 12 processors. Once the constant performance is reached, then adding additional processors to solve the problem provides no decrease in overall execution time. Upon examination of the 20-processor EABWulf results, the performance time begins to increase due to the increase in system overhead associated with additional processors. The occurrence of such an asymptotic performance point is a function of the type of problem and how parallelizable the solution can be designed and implemented. It is important to know where this asymptotic point is relative to the application software to minimize system cost.

The multiprocessor SGI computers (South, Norm, and Whitcomb) depict similar performance trends as the aforementioned clusters, that is, the performance increases (wall clock time decreases) with an increasing number of processors until the asymptotic performance/overhead barrier is reached. Based upon the results from Whitcomb and EABWulf it is speculated that performance gains on the order of

50 percent could be achieved by adding four additional processors to South, Norm, and GeoWulf, making them 12 processor machines.

It is difficult to perform an in-depth interpretation of a heterogeneous cluster such as BWolf. The value of performing an analysis using all eight processors is to demonstrate the potential computational resource of a PoPC assembled from discarded Pentium II and III computers. With eight processors BWolf's performance was 17 percent faster than the single processor performance of South.

The Opteron McCoy, run in a 32-bit single-processor mode, was approximately 70 percent slower than the single processor performance of the machine South, and 13 percent slower than Norm, as shown in figure 5. Since all the SGI machines were run in a 64-bit executable mode, a performance deficit was anticipated. However, McCoy's performance in 32-bit executable mode with two processors was only 17 percent slower than South's two-processor performance and 20 percent faster than Norm's two-processor performance. It is speculated that the dual-processor performance gain is due to improvements in multiprocessor hardware of the new Opteron workstations. When McCoy was run in a dual-processor 32-bit executable mode, linear scaling in performance was achieved with an increase in performance of 98 percent, see figure 6. No other computer evaluated achieved as large an increase in performance from the addition of one processor. GeoWulf exhibited the next highest increase in performance with 80 percent. While in contrast a dual processor case on South achieved only a 35-percent increase in performance.

McCoy run in 64-bit single-processor mode was 130 percent faster than South's single-processor performance. Since 64-bit multiprocessor runs on McCoy or Eureka were not possible due to compiler problems, there are no data for these conditions. However, based upon data trends from McCoy in 32-bit dual processor mode it is reasonable to speculate the Opteron performance over the SGI's would continue with increasing number of processors. Furthermore, with two processors in 64-bit executable mode, a potential 250-percent performance increase is realizable relative to South's two-processor performance.

If an Opteron processor performance scales linearly with processor speed, a reasonable assumption, then a 13-percent increase in performance is achievable from the workstation Eureka relative to McCoy in the two processor.

There is a considerable difference in the 32- to 64-bit performance results between the two benchmark programs. The results from the floating-point intensive HPL software showed little difference in performance between the 32- and 64-bit operations. Whereas using CFL3D there was almost a factor of four difference in performance. The contradictory nature of these results clearly shows the importance of including relevant application programs as part of one's benchmarking efforts.

The desktop dual processor computer Riker exhibited performance that is consistent with differences in clock speed between single processor results from EABWulf and GeoWulf. In dual-processor operation, the increase in performance was comparable to that of GeoWulf; this trend is consistent with the performance scaling of the Opteron McCoy computer in dual-processor mode. South's results in single processor mode was approximately 150 percent faster than the single processor Riker results but this performance difference substantially narrowed to only 87 percent faster in dual-processor mode. In comparing the results from the Opteron workstations, the clusters, and the desktop PC, it is obvious that incorporating a second processor into the system could be very performance and cost effective.

The Influence of System Cost

In any comparison of computers, some analysis of cost must be included and a reasonable metric is performance to cost ratio. Performance to cost ratios were calculated for each computer using the CFL3D performance results and the new equipment cost. Results for EABWulf and Bwulf are also presented using the additional expenditure costs associated with the assembly of these clusters along with original hardware costs. These results were normalized relative to the single processor performance to cost ratio of the computer system Whitcomb, and are presented in figure 7.

Another cost that is pertinent to mainframe computers is associated with annual maintenance agreements. Although new mainframe computers are frequently purchased with warranties, a rather substantial annual cost can be incurred with the extension of their warranties. Annual warranty costs are a function of the complexity of the mainframe and its age. For this study an annual cost of \$30K will be assumed.

Most commodity computers, even the Opteron workstations discussed herein, are not purchased with extended warranties. Many computers have 3- to 12-month warranties and if the computer has a total system failure after the warranty expires, the machine is frequently disposed of, instead of being repaired. The use of the annual maintenance cost associated with a mainframe computer to purchase new commodity computers is also discussed.

The SGI mainframe computers will be used as the basis of comparison. The fastest SGI machine, South, had the highest performance to cost ratio for the SGI computers and had approximately 5.5 times the performance to cost ratio of Whitcomb.

The new Opteron workstations exhibit considerable computational performance whether in 32- or 64-bit operation. Performance gains relative to the SGI's are attributed to both processor speed and system architecture. The performance to cost ratio of these machines is in excess of 380 times that of Whitcomb, and 60 times that of South, as seen in figure 7. This trend can be further evidenced in table 2, which represents the data used to create the curves in figure 7. Depending upon the configuration of dual-processor Opteron workstations four or more could be purchased annually for the same cost as the annual maintenance fee of the mainframe computer. The equivalent computing performance is approximately that of two eight-processor South computers.

Clusters based upon PoPCs provide comparable performance as multiprocessor SGI's. In this study GeoWulf's performance was comparable to Whitcomb and Norm. GeoWulf was purchased as a cluster and cost approximately \$20K whereas Whitcomb and Norm cost approximately \$300K new. GeoWulf's performance to cost ratio as compared to the SGI mainframe South is shown in figure 8. The first curve (diamond symbols) in figure 8 represents the total performance to cost of GeoWulf relative to a single processor of South. The second curve (square symbols) in figure 8 represents the total performance to cost ratio of GeoWulf per processor of South, where the performance ratio of GeoWulf at four processors is compared to that of South at four processors. On average, GeoWulf exhibits nearly twice the performance to cost ratio of South. Furthermore, the cost of GeoWulf is less than the annual maintenance fee of a mainframe computer. This means for the equivalent expenditure associated one can obtain equivalent computational capability as Whitcomb or Norm. Another benefit of using PoPC clusters is related to the cost to replace a failed processor. Replacement of a single node costs approximately 0.10 of the maintenance fee of a mainframe machine.

Desktop computers are frequently retired after 3 years of use. If one ignores the original equipment cost and just considers the cost of the additional expenditures necessary to create a cluster, these retired machines can provide considerable computational resources for an organization. Bwolf, the slowest of the three PC clusters, produced better computational performance than a single processor on South, two processors on Norm, and three processors on Whitcomb, as shown in figure 5. Similarly, the 20-processor cluster, EABWulf, had an equivalent processing power as three of South's processors, four processors of Norm, and six processors of Whitcomb. EABWulf's performance to cost ratio, based upon original equipment costs, is approximately 2.5 times that of Whitcomb. If this ratio is based upon new expenditure cost the performance to cost ratio is 50 times that of Whitcomb and nearly 10 times that of South, as seen in figure 7. The \$2500 to \$5000 cost to assemble BWolf or EABWulf is minimal compared to a fraction of the original equipment cost or even the annual maintenance fee of the mainframe machine. As nodes on PoPC clusters eventually fail, other desktop computers within the organization are almost continuously available as replacements, to further grow the cluster or develop new clusters. As long as the effects of heterogeneous cluster environments are taken into account for a given problem type, PoPC clusters can be a valuable computing asset.

Conclusions

A study of computer performance was conducted using legacy mainframes, state of the art workstations, PC based clusters, and desktop computers. Two benchmark programs were used: a synthetic and an application program. This study also assessed the influence of original equipment and annual maintenance cost.

The Opteron based computers achieved superior performance on a per processor basis. Performance gains were achieved through processor speed improvements and system architecture. When the cost of these Opteron machines was considered, their performance to price ratio is several multiples of the legacy mainframe computers.

Clusters based on PoPC exhibited comparable performance to the legacy mainframe computers. The major advantage of the clusters was their cost relative to that of the legacy machines. In the cases where retired PCs were used to form a cluster (EABWulf and Bwolf) the additional funds were minimal, less than \$5000. In a large organization where desktop computers are typically retired before they fail, these retired computers become a valuable reservoir of future computational resources. Even clusters purchased as new computers, such as GeoWulf, exhibited excellent performance to price relative to the legacy mainframe computers. Relative to the annual maintenance cost of mainframe computers it is possible to annually purchase equivalent or better PoPC clusters and computers.

The performance of single- and dual-processor desktop computers can rival the single processor performance of traditional mainframe computers. The cost of these machines is frequently less than \$3000 and provides an excellent resource for limited computing.

References

- ¹Moore, Gordon E.: Cramming More Components Onto Integrated Circuits. *Electronics*, vol. 38, no. 8, 1965.
- ²Top 500 Supercomputer sites, Top 500 Statistics November 200, <http://www.top500.org/sublist/stats/index.php?list=2003-11-30&type=archtype&submit=1> Accessed May 27, 2004.
- ³Sitaker, Kragen; Brown, Robert G.; and Lindahl, Greg: Beowulf Mailing List FAQ, Version 2, Rev May 1999. <http://www.canonical.org/~kragen/beowulf-faq.txt> Accessed May 24, 2004.
- ⁴Ridge, Daniel; Becker, Donald; Merkey, Phillip; and Sterling, Thomas: *Beowulf: Harnessing the Power of Parallelism in a Pile-of-PCs*, Paper from IEEE Aerospace 1997 Proceedings, 1997.
- ⁵Dongarra, Jack: Netlib Linpack Performance data. Rev May, 2004 <http://performance.netlib.org/performance/html/linpack.data.col0.html> Accessed May 27, 2004.
- ⁶Dongarra, Jack: Netlib Linpack Performance data. Rev May, 2004 <http://performance.netlib.org/performance/html/linpack.data.col0.html> Accessed May 27, 2004.
- ⁷Petit, Antoine; Whaley, R. Clint; Dongarra, Jack; and Cleary, Andrew J.: *HPL - A Portable Implementation of the High-Performance Linpack Benchmark for Distributed-Memory Computers*. Rev January 20, 2004. <http://www.netlib.org/benchmark/hpl/> Accessed May 24, 2004.
- ⁸Top 500 Supercomputer sites, *The Linpack Benchmark*, <http://www.top500.org/lists/linpack.php> Accessed May 24, 2004.
- ⁹HPL Frequently Asked Questions: "What process grid ratio $P \times Q$ should I use?" <http://www.netlib.org/benchmark/hpl/faqs.html> Accessed September 10, 2004.
- ¹⁰Biedron, Robert T.; and Rumsey, Christopher L.: *CFL3D Version 6*, Rev October 7, 2003. <http://cfl3d/Cfl3dv6.larc.nasa.gov/cfl3dv6.html> Accessed May 24, 2004.
- ¹¹Slater, John W.: *NPARC Alliance Validation Archive, ONERA M6 Wing*. <http://www.grc.nasa.gov/WWW/wind/valid/m6wing/m6wing.html> Accessed May 29, 2004.
- ¹²Biedron, Robert T.; and Rumsey, Christopher L.: *CFL3D ONERA M6 Wing*, Rev September 12, 2003. http://cfl3d.larc.nasa.gov/Cfl3dv6/cfl3dv6_testcases.html Accessed May 24, 2004.
- ¹³MSC Software Corporation, home of NASTRAN: <http://www.mscsoftware.com/> Accessed May 29, 2004.
- ¹⁴HKS Software, home of ABAQUS: <http://www.hks.com/> Accessed May 29, 2004.
- ¹⁵The MIPS® architecture is developed by MIPS Technologies, Inc.
- ¹⁶MIPS R14000 Processor used in the SGI Origin 300/3000 series of systems. The R14000 is an extension of the R10000 family of processors. Introduced by SGI in 2001.
- ¹⁷MIPS R12000 Processor, used in the SGI Origin 200/2000 and the SGI Onyx 2 systems. Introduced in 1998, as an extension of the R10000 family.
- ¹⁸MIPS R10000 Processor, introduced in 1995 by SGI. The R10000's are 4-issue, processors with 5 functional units. The R10k was used in the early SGI Origin 200/2000 and SGI Onyx and SGI Onyx 2 systems.

¹⁹Trustees of Indiana University, *LAM/MPI Overview*, Rev Mar 14, 2004. <http://www.lam-mpi.org/about/overview/> Accessed May 24, 2004.

²⁰Indiana University: Pervasive Technology Labs. <http://www.pervasive.iu.edu/> Accessed May 27, 2004.

²¹Ohio Supercomputer Center, <http://www.osc.edu/> Accessed May 27, 2004.

²²*HPL Frequently Asked Questions: "What process grid ratio $P \times Q$ should I use?"*
<http://www.netlib.org/benchmark/hpl/faqs.html> Accessed September 10, 2004.

²³Linux Scalability Effort, *NUMA Frequently Asked Questions*, Rev, January 4, 2002.
<http://lse.sourceforge.net/numa/faq/> Accessed May 31, 2004.

Table 1. Table of Used Computational Hardware, Showing Icons Used, # Processors, Type of CPU and Architecture















Family	Name	Icon(s) used	# Processors	CPU	CPU Arch
SGI:	Whitcomb		16	250Mhz	MIPS R10k
	Norm		8	400Mhz	MIPS R12k
	South		8	600Mhz	MIPS R14k
Cluster:	BWolf	 Total Cost	8	450-700Mhz	Pentium 3
		 New Cost			
	EABWulf	 Total Cost	20 (SMP)	833Mhz	Pentium 3 Xeon
		 New Cost			
	GeoWulf		8	2.4Ghz	Pentium 4
PC:	Riker		2 (SMP)	1.7Ghz	Athlon MP 1900
	McCoy	 32 bit	2 (SMP)	1.6Ghz	Opteron 242
		 64 Bit			
		 64 Bit (Figure 2)			
	Eureka	 32 bit (Figure 2)	4 (SMP)	1.8Ghz	Opteron 844
		 64 Bit (Figure 2)			

Table 2. CFL3D Runtime Performance/Cost Table

System Name	No. of Processors In system	System Cost (\$)	Uniprocessor Performance (Sec.)	Multiprocessor Performance (Sec.)	Performance/Cost Normalized ¹	Performance /Cost Normalized ²
Bwolf (New Cost)	8	2500	NA	2.542	NA	304.99
Bwolf (Total Cost)	8	20000	NA	2.542	NA	38.124
EABWulf (New Cost)	20	5000	1520.203	5.361	33.075	358.825
EABWulf (Total Cost)	20	65000	1520.203	5.361	2.544	24.741
GeoWulf	8	20000	777.820	329.708	16.161	98.558
McCoy (32-bit)	2	4000	693.210	329.708	91.663	179.315
McCoy (64-bit)	1	4000	165.297	156.326	380.224	380.224
Norm	8	300000	604.016	156.326	1.387	9.059
Riker	2	2100	1041.312	127.539	114.965	213.252
South	8	110000	303.000	350.500	5.584	43.121
Whitcomb	16	300000	838.000	165.297	1.000	13.567

1 – One Processor performance and system cost normalized to 1 Processor Whitcomb

2 – All Processors performance and system cost normalized to 1 Processor Whitcomb

```

HPLinpack benchmark input file
Innovative Computing Laboratory, University of Tennessee
HPL.out    output file name (if any)
6          device out (6=stdout,7=stderr,file)
14         # of problems sizes (N)
1000 2000 3000 4000 5000 6000 7000 8000 9000 10000 12000 15000 17000 20000 Ns  ## MEMORY LIMITED ##
13         # of NBs
16 30 32 38 40 48 64 80 100 120 128 200 256  NBs
3          # of process grids (P x Q)  ## VARIABLE ##
1 1 2     Ps          ## VARIABLE ##
1 2 1     Qs          ## VARIABLE ##
16.0      threshold
1          # of panel fact
2 1 2     PFACTs (0=left, 1=Crout, 2=Right)
1          # of recursive stopping criterium
8 4       NBMINs (>= 1)
1          # of panels in recursion
2         NDIVs
1          # of recursive panel fact.
2 1 2     RFACTs (0=left, 1=Crout, 2=Right)
2          # of broadcast
1 3       BCASTs (0=1rg,1=1rM,2=2rg,3=2rM,4= Lng,5=LnM)
1          # of lookahead depth
1         DEPTHs (>=0)
2         SWAP (0=bin-exch,1=long,2=mix)
64        swapping threshold
0         L1 in (0=transposed,1=no-transposed) form
0         U  in (0=transposed,1=no-transposed) form
1         Equilibration (0=no,1=yes)
8         memory alignment in double (> 0)

```

Figure 1. Sample HPL.dat file.

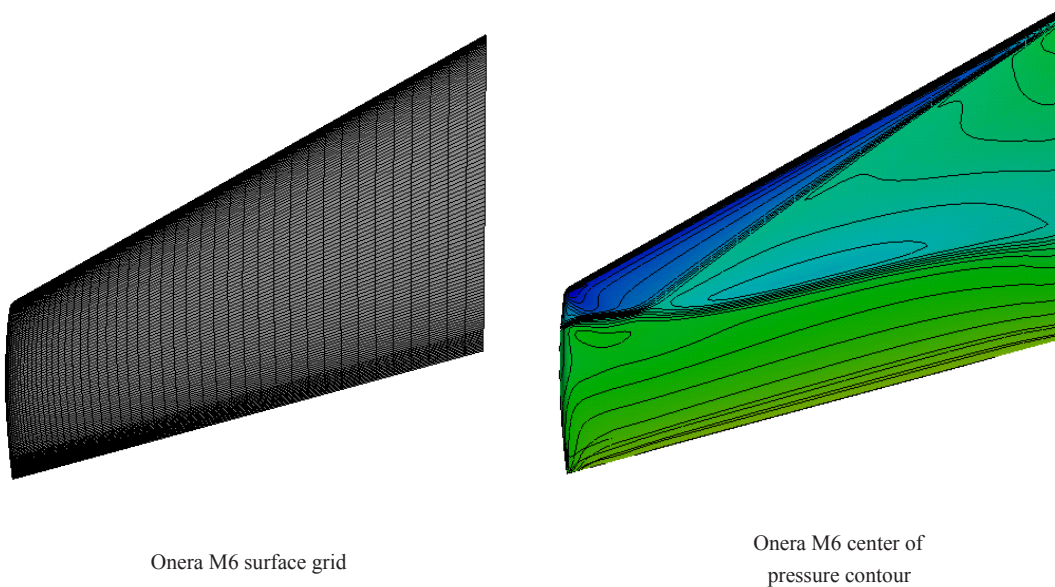


Figure 2. Onera M6 grid and pressure contours.

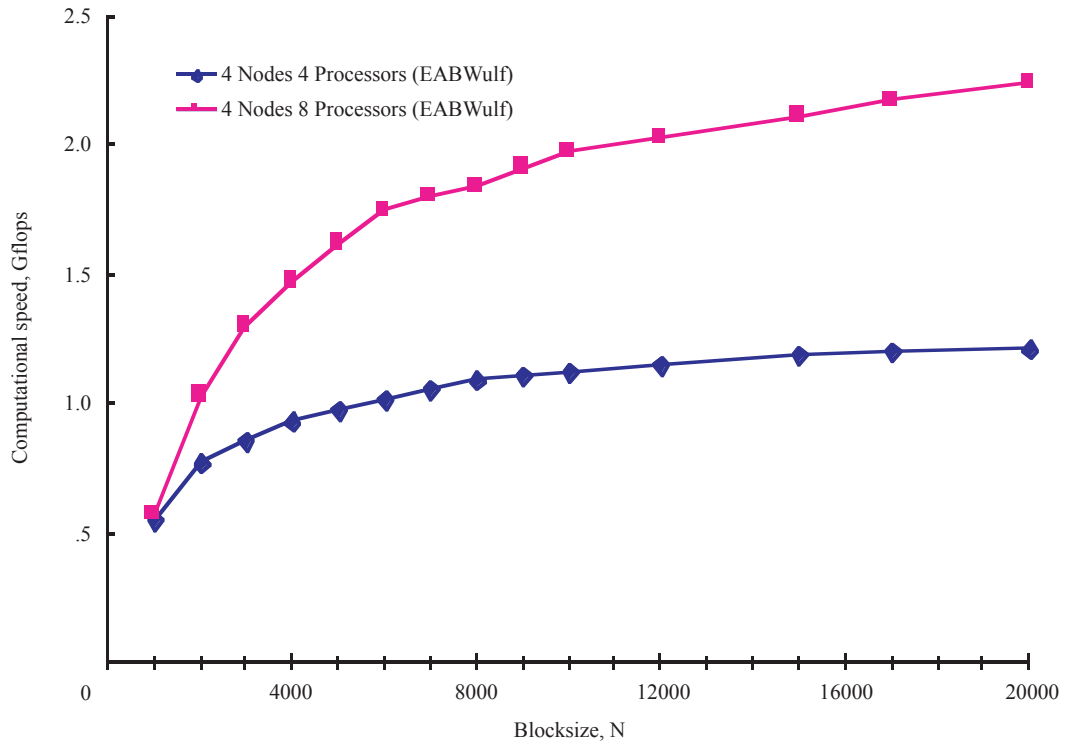


Figure 3. Four processor versus eight processor peak HPL computational performance on EABWulf.

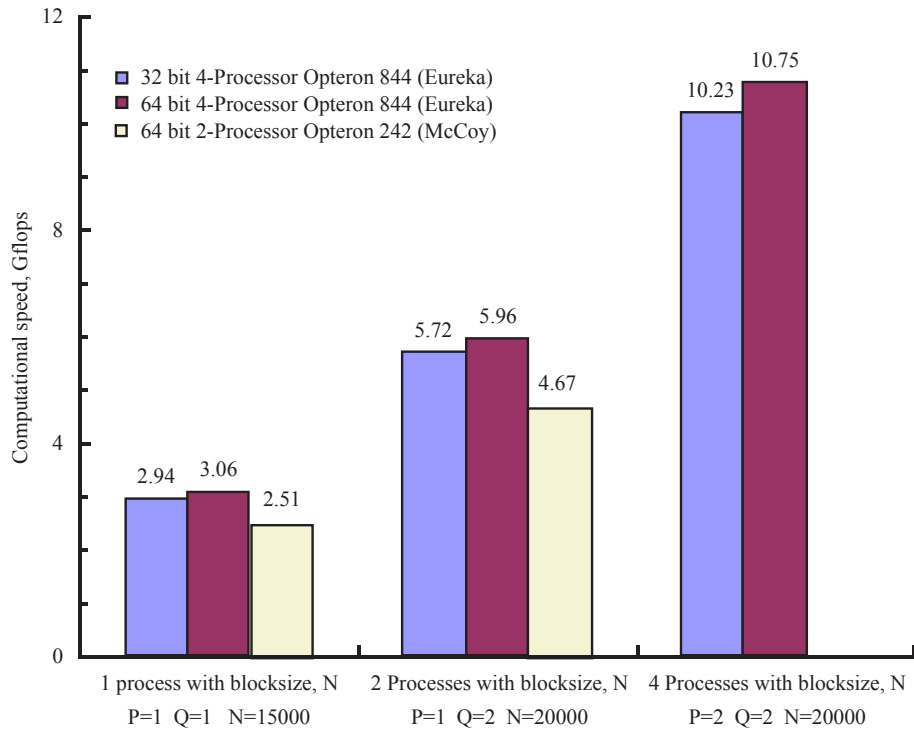


Figure 4. Peak Operton HPL performance for varying matrix (P x Q) configurations, and blocksize (N).

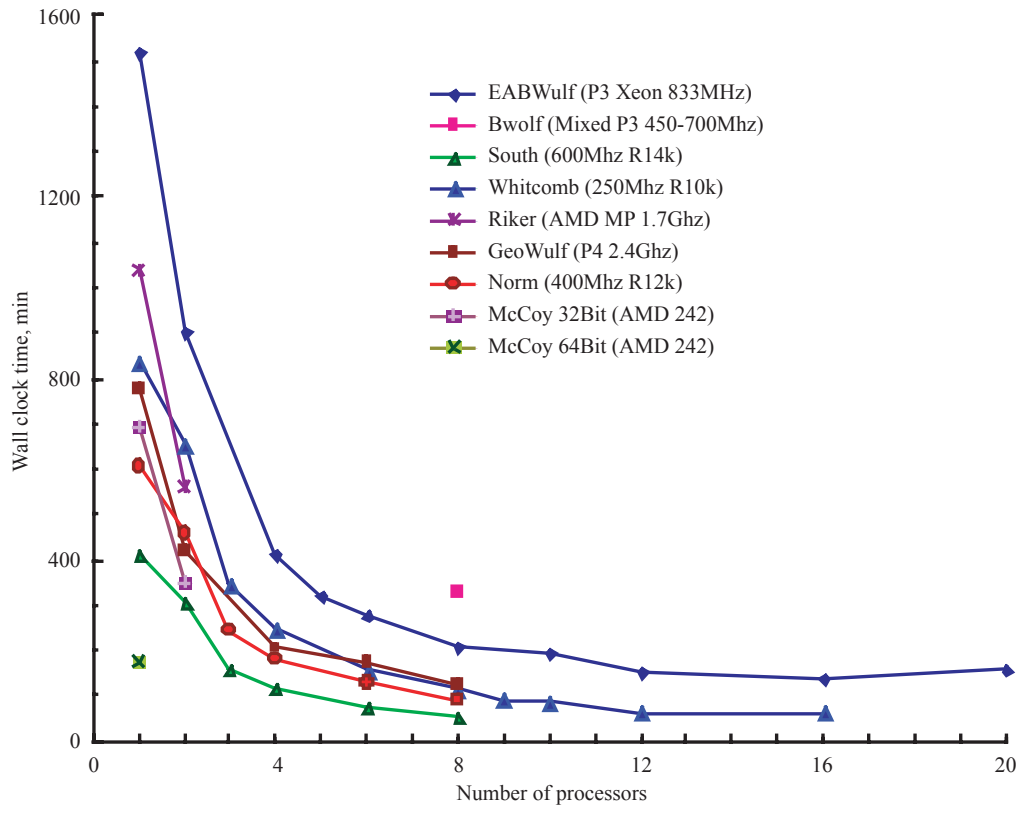


Figure 5. Runtime performance of CFL3D using the Onera M6 wing case.

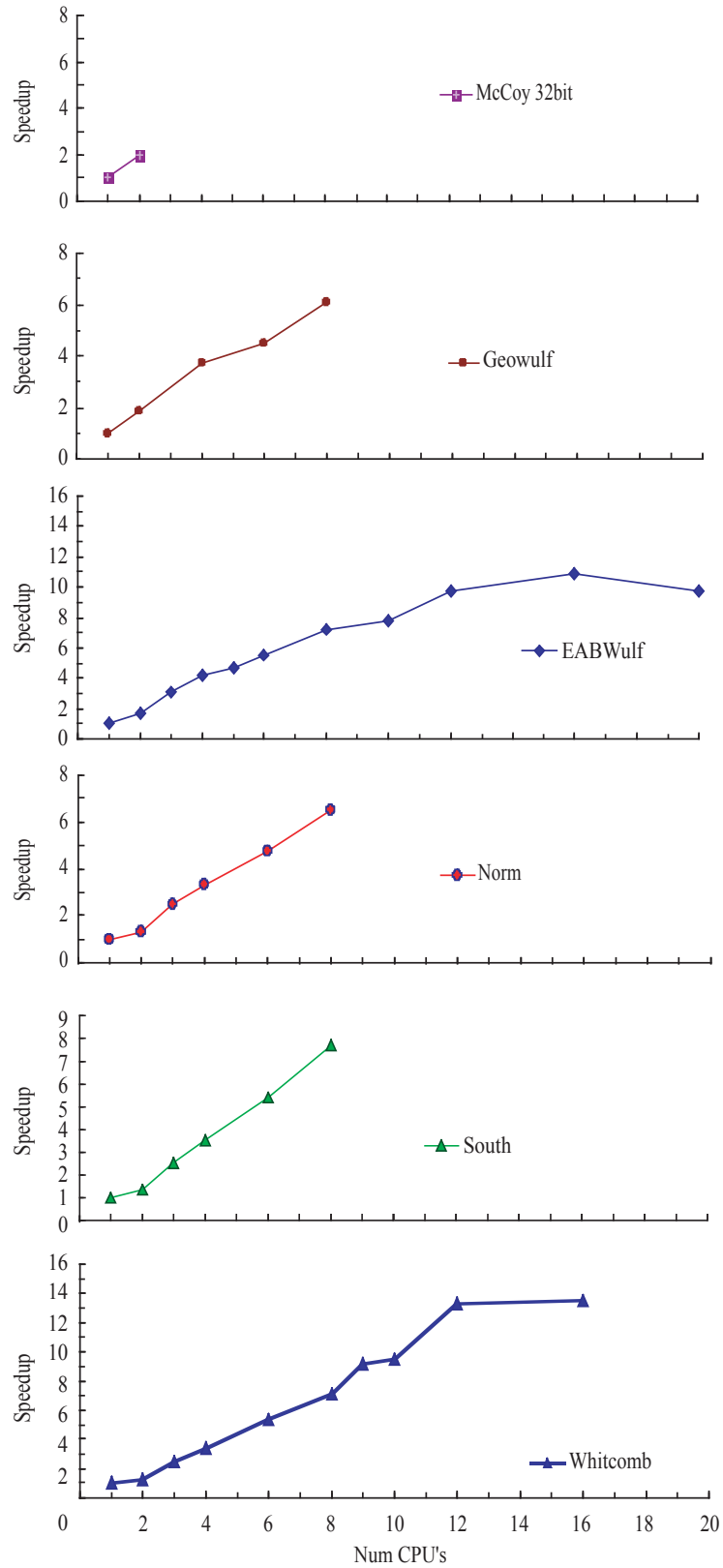


Figure 6. CFL3D multiprocessor speedups, normalized by one processor of the respective machine.

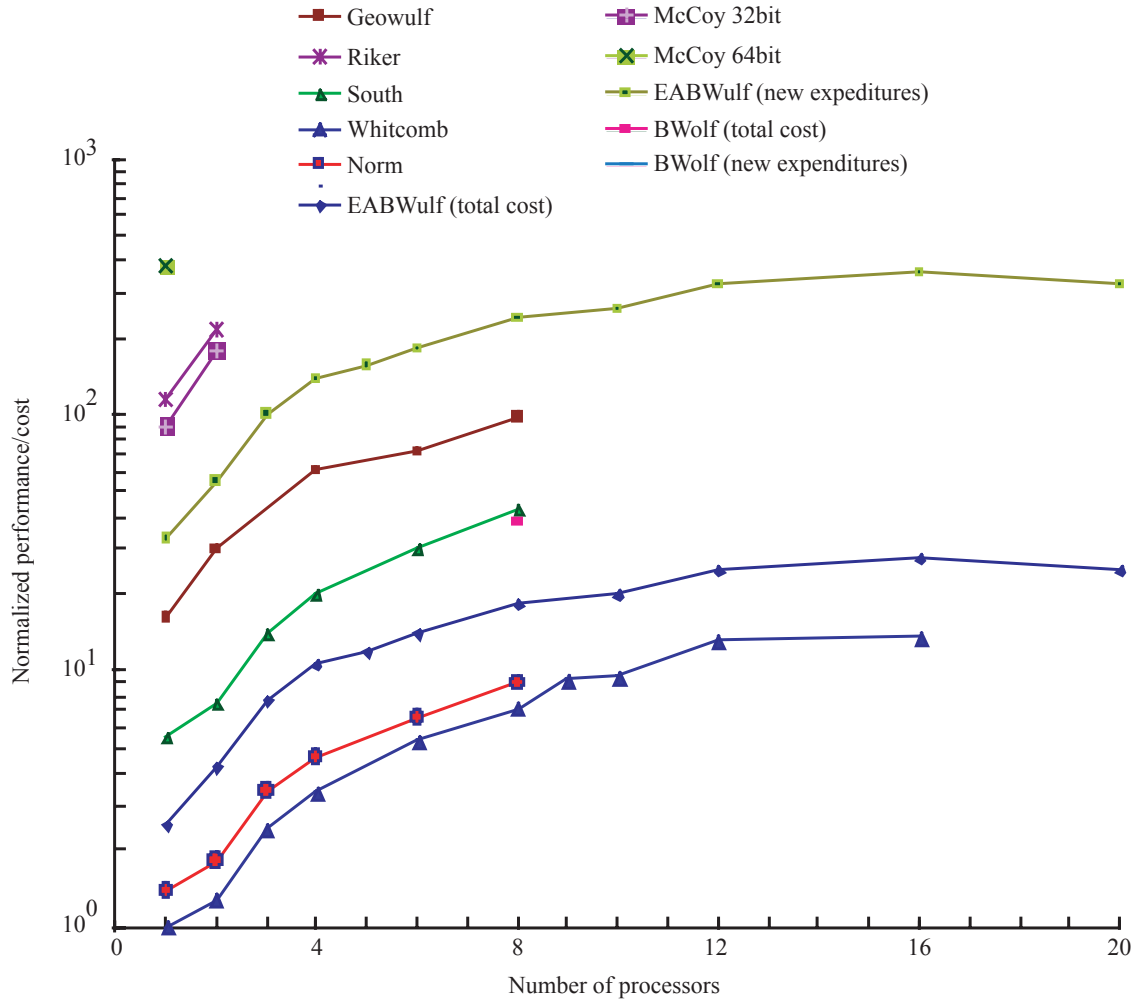


Figure 7. CFL3D runtime performance/cost normalized to one processor on Whitcomb versus the total number of processors from the respective hardware.

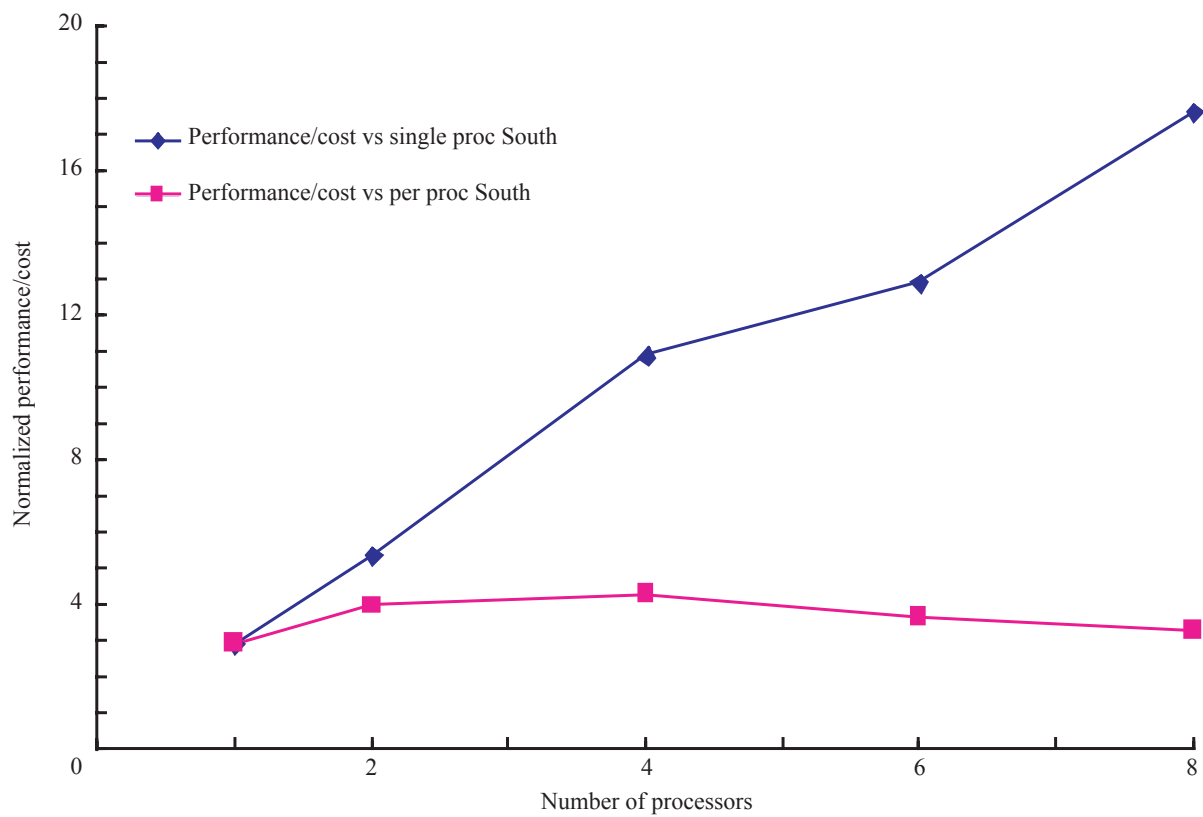


Figure 8. CFL3D performance/cost of GeoWulf versus South on a per-processor and single-processor case.

REPORT DOCUMENTATION PAGE					Form Approved OMB No. 0704-0188	
<p>The public reporting burden for this collection of information is estimated to average 1 hour per response, including the time for reviewing instructions, searching existing data sources, gathering and maintaining the data needed, and completing and reviewing the collection of information. Send comments regarding this burden estimate or any other aspect of this collection of information, including suggestions for reducing this burden, to Department of Defense, Washington Headquarters Services, Directorate for Information Operations and Reports (0704-0188), 1215 Jefferson Davis Highway, Suite 1204, Arlington, VA 22202-4302. Respondents should be aware that notwithstanding any other provision of law, no person shall be subject to any penalty for failing to comply with a collection of information if it does not display a currently valid OMB control number.</p> <p>PLEASE DO NOT RETURN YOUR FORM TO THE ABOVE ADDRESS.</p>						
1. REPORT DATE (DD-MM-YYYY)		2. REPORT TYPE			3. DATES COVERED (From - To)	
01- 01 - 2005		Technical Memorandum				
4. TITLE AND SUBTITLE Performance Comparison of Mainframe, Workstations, Clusters, and Desktop Computers				5a. CONTRACT NUMBER		
				5b. GRANT NUMBER		
				5c. PROGRAM ELEMENT NUMBER		
6. AUTHOR(S) Farley, Douglas L.				5d. PROJECT NUMBER		
				5e. TASK NUMBER		
				5f. WORK UNIT NUMBER 23-728-30-11		
7. PERFORMING ORGANIZATION NAME(S) AND ADDRESS(ES) NASA Langley Research Center Hampton, VA 23681-2199				8. PERFORMING ORGANIZATION REPORT NUMBER L-18392		
9. SPONSORING/MONITORING AGENCY NAME(S) AND ADDRESS(ES) National Aeronautics and Space Administration Washington, DC 20546-0001				10. SPONSOR/MONITOR'S ACRONYM(S) NASA		
				11. SPONSOR/MONITOR'S REPORT NUMBER(S) NASA/TM-2005-213505		
12. DISTRIBUTION/AVAILABILITY STATEMENT Unclassified - Unlimited Subject Category 62 Availability: NASA CASI (301) 621-0390						
13. SUPPLEMENTARY NOTES Farley, Langley Research Center, Hampton, VA. An electronic version can be found at http://ntrs.nasa.gov						
14. ABSTRACT A performance evaluation of a variety of computers frequently found in a scientific or engineering research environment was conducted using a synthetic and application program benchmarks. From a performance perspective, emerging commodity processors have superior performance relative to legacy mainframe computers. In many cases, the PC clusters exhibited comparable performance with traditional mainframe hardware when 8-12 processors were used. The main advantage of the PC clusters was related to their cost. Regardless of whether the clusters were built from new computers or whether they were created from retired computers their performance to cost ratio was superior to the legacy mainframe computers. Finally, the typical annual maintenance cost of legacy mainframe computers is several times the cost of new equipment such as multiprocessor PC workstations. The savings from eliminating the annual maintenance fee on legacy hardware can result in a yearly increase in total computational capability for an organization.						
15. SUBJECT TERMS Beowulf clusters; Parallel computers; Computer performance; Cluster design; Computer design; POPC						
16. SECURITY CLASSIFICATION OF:			17. LIMITATION OF ABSTRACT	18. NUMBER OF PAGES	19a. NAME OF RESPONSIBLE PERSON	
a. REPORT	b. ABSTRACT	c. THIS PAGE			STI Help Desk (email: help@sti.nasa.gov)	
U	U	U	UU	26	19b. TELEPHONE NUMBER (Include area code) (301) 621-0390	